



Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Short communication

The effect of the stability threshold on time to stabilization and its reliability following a single leg drop jump landing

Duncan P. Fransz^{a,b,*}, Arnold Huurnink^{a,c}, Vosse A. de Boode^d, Idsart Kingma^a,
Jaap H. van Dieën^a

^a MOVE Research Institute Amsterdam, Department of Human Movement Sciences, VU University, Amsterdam, The Netherlands^b Department of Orthopaedic Surgery, Atrium Medical Center, Heerlen, The Netherlands^c Department of Nuclear Medicine, Academic Medical Center, Amsterdam, The Netherlands^d Adidas miCoach Performance Centre, AFC Ajax, Amsterdam, The Netherlands

ARTICLE INFO

Article history:

Accepted 28 December 2015

Keywords:

Force plate
Force platform
Outcome measure
Dynamic
Postural control

ABSTRACT

We aimed to provide insight in how threshold selection affects time to stabilization (TTS) and its reliability to support selection of methods to determine TTS.

Eighty-two elite youth soccer players performed six single leg drop jump landings. The TTS was calculated based on four processed signals: raw ground reaction force (GRF) signal (RAW), moving root mean square window (RMS), sequential average (SA) or unbounded third order polynomial fit (TOP). For each trial and processing method a wide range of thresholds was applied. Per threshold, reliability of the TTS was assessed through intra-class correlation coefficients (ICC) for the vertical (V), anteroposterior (AP) and mediolateral (ML) direction of force.

Low thresholds resulted in a sharp increase of TTS values and in the percentage of trials in which TTS exceeded trial duration. The TTS and ICC were essentially similar for RAW and RMS in all directions; ICC's were mostly 'insufficient' (< 0.4) to 'fair' (0.4–0.6) for the entire range of thresholds. The SA signals resulted in the most stable ICC values across thresholds, being 'substantial' (> 0.8) for V, and 'moderate' (0.6–0.8) for AP and ML. The ICC's for TOP were 'substantial' for V, 'moderate' for AP, and 'fair' for ML.

The present findings did not reveal an optimal threshold to assess TTS in elite youth soccer players following a single leg drop jump landing. Irrespective of threshold selection, the SA and TOP methods yielded sufficiently reliable TTS values, while for RAW and RMS the reliability was insufficient to differentiate between players.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

A variety of laboratory and field tests have been applied to characterize different aspects of a soccer player's ability as potential determinants of performance (Svensson and Drust, 2005). Among these aspects dynamic postural stability has received increasing interest (Fransz et al., 2013, Shaw et al., 2008). The most commonly applied dynamic postural stability test is the single leg jump landing, which involves subjects jumping either from a box or to a certain height, landing upon a force plate on one foot, and stabilizing as quickly as possible. The performance on such a test is usually quantified with the ground reaction forces

(GRF) from which a variety of outcome measures can be calculated (Fransz et al., 2013).

Time to stabilization (TTS) is a frequently used outcome measure, but widely different calculation methods to establish TTS following a single leg jump landing test have been identified in the literature (Fransz et al., 2015). For instance, these calculation methods differ with regard to the input signal that is used. The GRF's can be assessed in vertical (V), anteroposterior (AP) or mediolateral (ML) direction, resulting in substantially different TTS values, ranging from 1.3 to 6.1 s (Fransz et al., 2015). Secondly, four essentially different signal-processing methods have been employed. The 'RAW' method employed no additional processing (Colby et al., 1999), the 'RMS' method produced a new signal by calculating the root mean square over a time window moving along the time series (Tulloch et al., 2012), the 'SA' method produced a sequential average signal by calculating a new average after each added data point (Colby et al., 1999), and the 'TOP'

* Corresponding author at: MOVE Research Institute Amsterdam, Department of Human Movement Sciences, Van der Boechorststraat 9, 1081 BT, Amsterdam, The Netherlands. Tel: +3120 5988492; Fax: +3120 5988529.

E-mail address: dpfransz@gmail.com (D.P. Fransz).

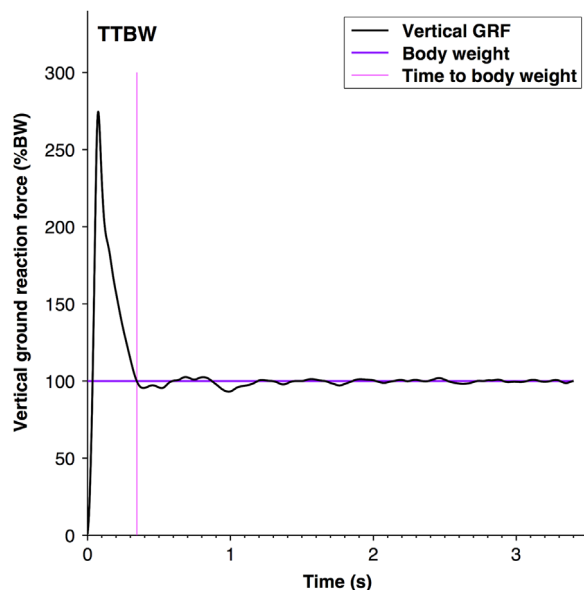


Fig. 1. A typical example of the establishment of 'time to bodyweight' (TTBW). The TTBW is the intersection of the Vertical GRF signal with the body weight (100%) directly following the impact peak.

method produced a new signal by fitting an unbounded third order polynomial to the time series (Ross and Guskiewicz, 2003). These processed signals resulted in a range of mean TTS values differing up to 5.5 s within the same direction (Fransz et al., 2015). Furthermore, a definition of the stable state has to be determined. This is the threshold below which the processed signal is considered stable. The effects of threshold selection have not yet been assessed. This is important as threshold selection may considerably change the resulting TTS values. Even more so, it may have implications for the interpretation of the values, as higher thresholds lead to smaller TTS values (Fransz et al., 2015).

To date it is unknown which processed signal/threshold combination will be most accurate in detecting impairments in sensorimotor function. A prerequisite for accurate detection is reliability and therefore a thorough and systematic examination of the effect of threshold selection in combination with the different processing methods seems warranted. The present study provides an analysis of reliability for a wide range of thresholds in combination with four commonly used signal processing methods and three directions of force.

2. Methods

2.1. Participants

At the youth academy of AFC Ajax, players are regularly tested as part of a program aiming to monitor player performance on a variety of neuromusculoskeletal tasks. The current performance data set was acquired at the start of the 2013–2014 season. We included the data of 82 players between 11 and 18 years old (mean \pm SD; age 14.10 ± 1.86 years; height 1.68 ± 0.12 m; body weight 56.70 ± 13.20 kg) for whom six valid trials obtained during one session were available. At the time of measurements, all players were fit to perform at the highest standard of competitive football matches. The local ethics committee approved the research protocol and all players or parents/guardians (depending on the age of the participant) were informed in advance of the procedures involved and provided written informed consent.

2.2. Instrumentation

Ground reaction forces (GRF) in vertical (V), anteroposterior (AP) and medio-lateral (ML) directions were recorded at 1000 samples/s, using a 40×60 cm AMTI force plate (type BP400600HF, Advanced Medical Technologies Inc., Watertown, MA, USA).

2.3. Procedures

The players were asked to jump from an aerobic step of 20 cm height, which was placed 5 cm posterior to the force plate. Players took off by means of a small jump with two feet, landed on the testing leg on the center of the force plate, and stabilized as quickly as possible. They had to balance for 15 s with their hands on their hips, whilst keeping all other movement to a minimum. If a player exaggerated the jump height or only slid of the aerobic step, he was asked to perform another jump following extra instruction. No specific instructions were given with regard to stabilization, however all players performed the test in a similar fashion. Following landing, the testing leg was flexed about 15° at the knee, the non-testing leg was flexed about 90° at the knee. Before actual testing commenced, all players completed the regular warm-up as accustomed before a training session and performed one practice trial per leg. Both legs were tested thrice without breaks; the left leg was appointed the initial testing leg. All trials were performed without shoes. A trial was considered invalid if a player touched the floor with the contralateral leg or if arm movement was used to regain balance.

2.4. Data processing

A custom MATLAB (The Mathworks, Natick, RI, USA) program was written for data processing. Raw GRF data were cropped from time of impact (vertical GRF > 10 N) to 12 s post-impact and rectified. Data were low pass filtered at 12 Hz with a bidirectional second order Butterworth filter (Huurnink et al., 2013).

2.5. Data analysis

To facilitate consistent data processing in V, AP and ML directions, the mean GRF value over 7 to 12 s was subtracted from the signal for each direction. Subsequently, the signals were rectified. Four processed signals were evaluated: (1) no additional processing (RAW) (Colby et al., 1999), (2) a new signal was produced by calculating the root mean square (RMS) over time windows of 250 ms, moving along the time series with 1 ms per step (Tulloch et al., 2012), (3) a sequential average (SA) signal was established by adding one data point at a time, and calculating a new average after each added point (Colby et al., 1999), and (4) an unbounded third order polynomial fit (TOP), which started at the peak GRF, was obtained by least-squares fitting the following function: $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, where $a_3 \neq 0$ (Ross and Guskiewicz, 2003).

Furthermore, a wide range of thresholds was applied. The SD over the last 5 s of each trial (7 to 12 s) was used to calculate up to 1600 threshold values, ranging from 0 to 160 SD (with steps of 0.1 SD). The TTS was defined as time between impact and the intersection of the processed signal with the threshold, after which it remained below the threshold for the subsequent 0.5 s (Tulloch et al., 2012). Consequently, TTS was calculated 1600 times for each trial. The effect of threshold level on the TTS and its reliability was assessed for twelve calculation methods (4 processed signals \times 3 GRF directions). For some trials the applied threshold did not result in a TTS value, i.e. when the processed signal stayed above the threshold for the duration of the entire trial. Therefore, the percentage of trials with no TTS value was calculated for each threshold per calculation method (% of 492 trials, 82 subjects \times 3 trials \times 2 legs).

Finally, in order to define the phase that primarily concerns the impact of the landing, we established the mean 'time to bodyweight' (TTBW) (based on 492 trials). This is the intersection of the RAW V signal with the body weight (100%) directly following the impact peak (Fig. 1). In further data processing, we ignored thresholds yielding TTS values shorter than mean TTBW.

2.6. Statistical analysis

The reliability of TTS across the six trials (3 trials \times 2 legs) per subject was calculated for each threshold and calculation method. An absolute agreement two-way random model (average measures) was applied to calculate intra-class correlation coefficients (ICC) (de Vet et al., 2006). When one or more trials did not reveal a TTS value, the ICC was discarded. Both limbs were grouped together, assuming that postural stability is predominately an indicator of whole body sensorimotor function (Witchalls et al., 2012). We used the considerations by Shrout (1998): 'insufficient' (< 0.40), 'fair' (0.40–0.60), 'moderate' (0.60–0.80), and substantial (> 0.80) (Shrout, 1998).

3. Results

The mean SDs of the GRF (in %BW) over the last 5 s of each trial (7 to 12 s) were 0.53 for vertical, 0.23 for anteroposterior and 0.27 for mediolateral GRF, respectively. Fig. 2 provides typical examples of the processed signals in relation to three different thresholds (i.e. 10, 30 and 50 SD) for each force direction. In contrast,

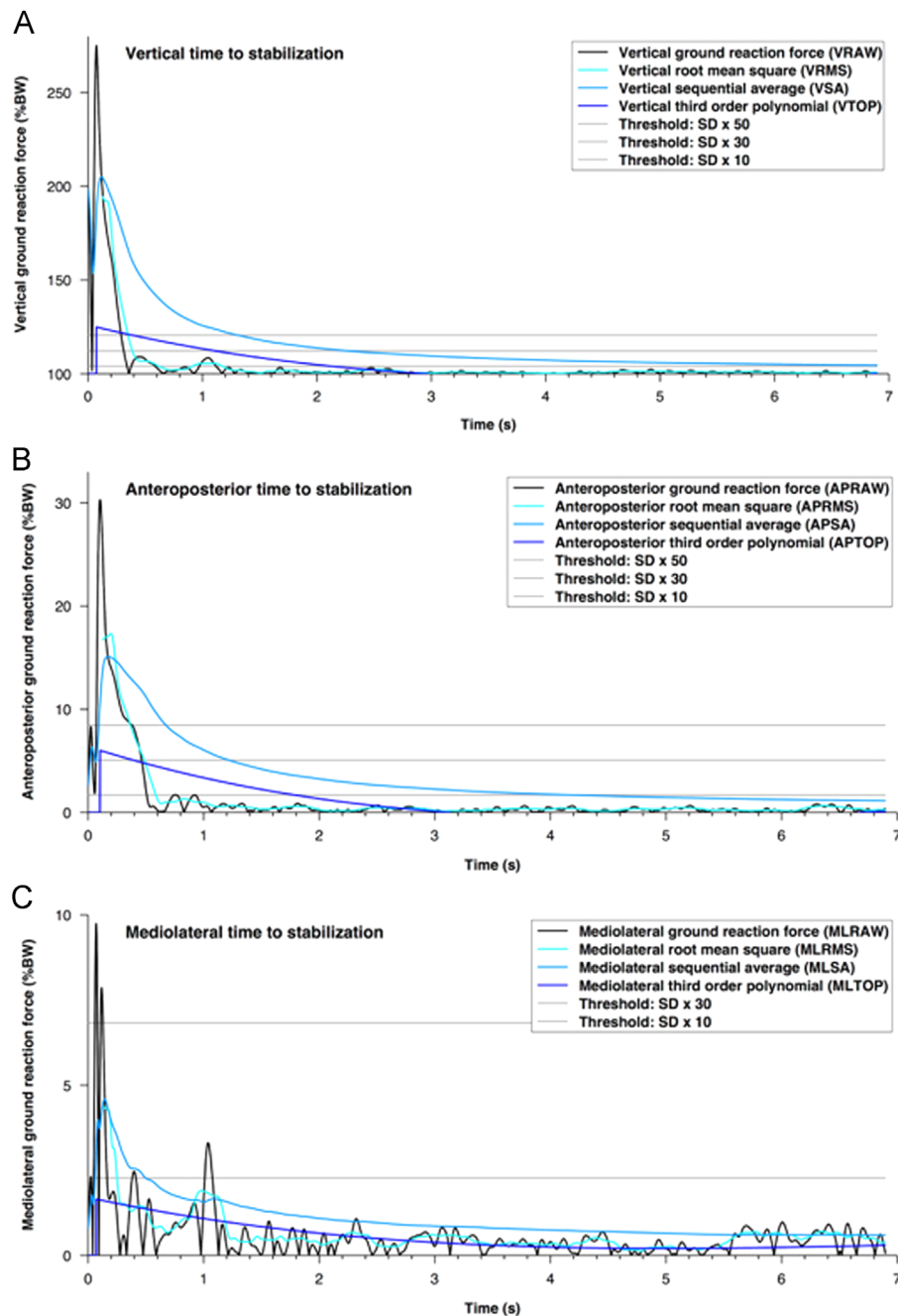


Fig. 2. Typical examples of the processed signals in relation to different thresholds (i.e. 10, 30 and 50 SD) for each direction of force (A. Vertical, B. Anteroposterior, and C. Mediolateral).

especially in V and AP directions, the SA and TOP signals stay at a higher level compared to RAW and RMS for a considerable time, which is caused by the relative high peak GRF.

Figs. 3 to 6 show how thresholds affect the mean TTS, the percentage of unusable trials and, when 100% of the trials resulted in a TTS, the ICC values. The end of the horizontal axes was set at the mean TTBW (0.30 ± 0.05 s), which varied across directions and methods from 11 SD (Fig. 6c) to 160 SD (Fig. 5a).

Overall, while most ICCs showed a clear drop at low threshold values, none of the combinations of method and direction showed a clear ICC peak, which would indicate an optimum threshold. The patterns of the TTS values and resultant ICC values were essentially similar for RAW and RMS, with ICC values being mostly 'insufficient' to 'fair' for the entire range of thresholds (Figs. 3 and

4). The ICC was consistently higher when the TTS approximated mean TTBW for all three directions. The SA signals resulted in the most stable ICC values across the threshold levels, being 'substantial' for V, and 'moderate' for AP and ML (Fig. 5). The ICC values for TOP were mostly 'substantial' for V, 'moderate' for AP, and 'fair' for ML (Fig. 6).

Table 1 shows an overview of previously reported mean TTS and ICC values.

4. Discussion

The main finding of the current study is that TTS is highly dependent on the threshold used, especially at lower threshold

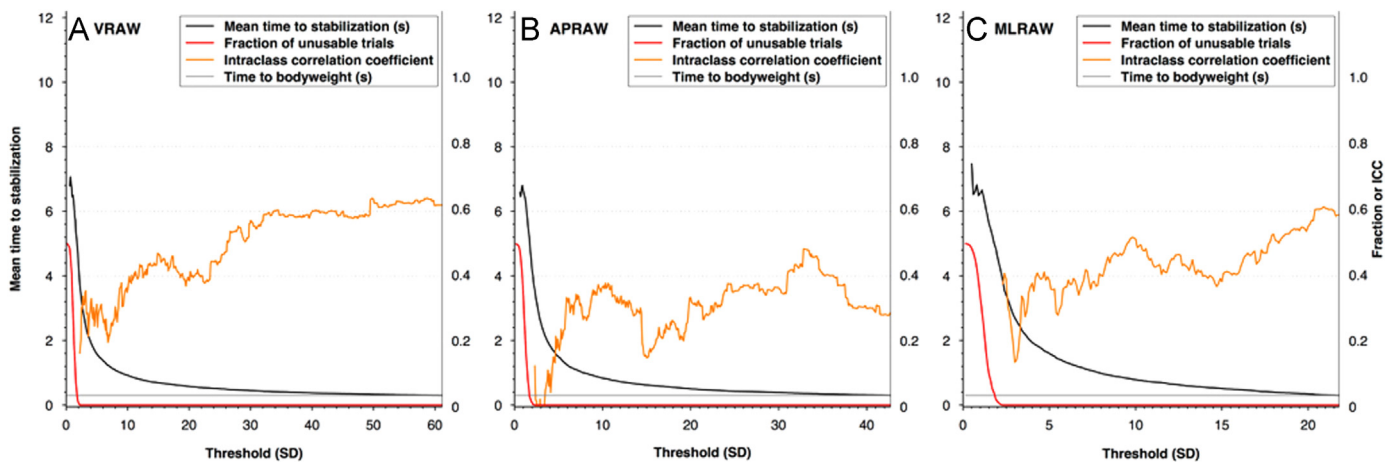


Fig. 3. The mean time to stabilization (TTS), the fraction of unusable trials, the intra class correlation (ICC) values and the mean time to body weight (TTBW) with regard to the RAW signals (A. Vertical, B. Anteroposterior, and C. Mediolateral). Note that the threshold magnitude along the x-axis differs per panel.

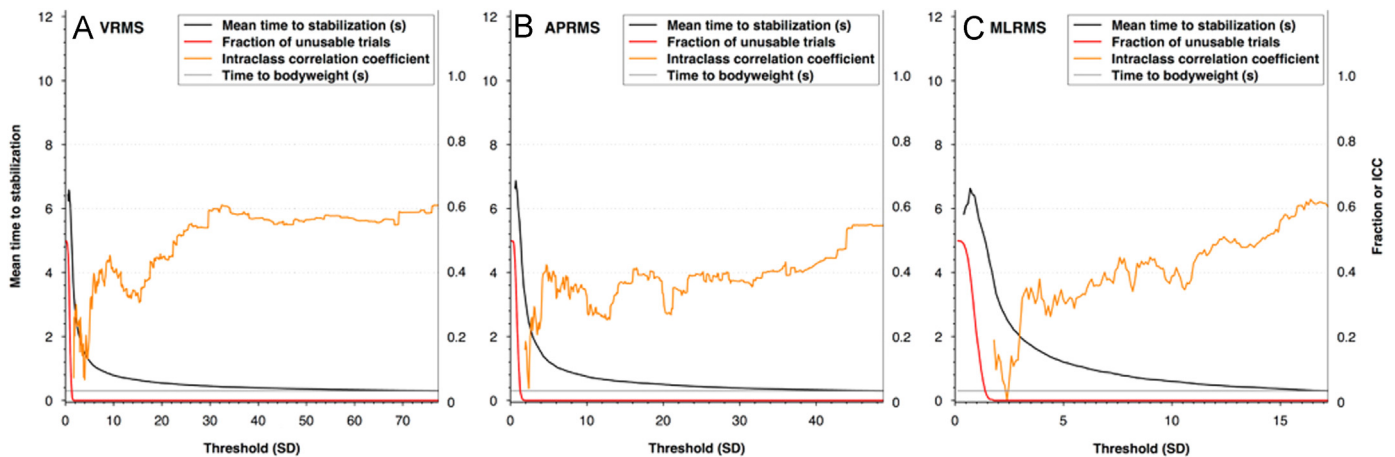


Fig. 4. Mean TTS, fraction of unusable trials, ICC values and mean TTBW with regard to the RMS signals (A. Vertical, B. Anteroposterior, and C. Mediolateral). Threshold magnitude along the x-axis differs per panel.

values. In addition, none of the processed signal/threshold combinations yielded a narrow band of threshold values with optimum reliability.

Regardless of threshold, RAW and RMS methods are not reliable enough to differentiate between subjects. A possible explanation could lie within incidental peaks of the GRF signal in the static phase. These methods are particularly sensitive to these random peaks, more likely to represent behavioral variance than a characteristic of individual performance.

In contrast, the SA and TOP methods yielded sufficiently reliable TTS values, provided that thresholds are not set too low. The threshold should be at least 30 SD for V, at least 15 SD for AP, and at least 7 SD for ML. Previous studies mostly determined the threshold per trial calculating the mean GRF over a certain time window, plus either 0.25 or 3 SD of the same time window (Table 1), though all of these calculations result in a horizontal line at a certain value. Some studies calculated the SD over the entire trial, including the impact peak, instead of the SD over the stable phase (7 to 12 s). In the present study this would have resulted in mean SDs of 13.9, 3.3, and 0.9 for V, AP, ML, respectively.

When comparing the current ICC values for the SA and TOP methods with previous research, our highest ICC values were in the lower range of previously reported values, at TTS values that were substantially lower (Table 1). The differences in ICC values

may be due to subject selection, as youth players may be less consistent in performance, a limited number of repetitions, and on site testing rather than laboratory conditions. The differences in TTS values may be related to our experimental setup, with a starting position lower and closer to the force plate than in previous studies. Furthermore, we did not control for jump height. This has most likely added additional variance compared to some previous studies (Brown et al., 2004, Ross et al., 2005, Wikstrom et al., 2005). However, this will most likely not result in lower reliability, as jump height will increase both within and between subject variance. In addition, in view of the high level of fitness and limited number of jumps, no breaks were given, assuming that possible fatigue would be negligible.

As the ICC values increased when the TTS approximated mean TTBW for the RAW, RMS and SA signals (in all three directions), it might be worthwhile to further examine this particular phase following a single leg drop jump landing. Possibly this could reveal new important characteristics of the force time series, in addition to the impact phase ('peak GRF' and 'time to peak GRF') (Fransz et al., 2013) and the static phase ('center of pressure (COP) speed', 'COP sway' and 'horizontal GRF') (Fransz et al., 2014), that may discriminate in stabilizing performance.

One should keep in mind that reliability is a prerequisite for sensitivity and thus usability in performance assessment or for

Table 1

An overview of previous studies regarding time to stabilization that reported intra class correlation (ICC) values. Mean TTS is in seconds.

Processed signal	Study	Threshold	n trials	Mean TTS	ICC average	ICC single
VRAW	Colby et al. (1999)	Body weight $\pm 5\%$	10	0.85	0.83	
	Wikstrom et al. (2005)	Body weight (during 5 s static stance) $\pm 5\%$	3	1.70		0.78
	Flanagan et al. (2008)	Body weight $\pm 5\%$	3	0.97	0.69	0.42
	Ebben et al. (2010)	Body weight $\pm 5\%$	3	0.85	0.64	
APRMS	Tulloch et al. (2012)	Minimal range for 10–15 s or 15–20 s window	2	3.76	0.43	
MLRMS	Tulloch et al. (2012)	Minimal range for 10–15 s or 15–20 s window	2	4.10	0.51	
VSA	Colby et al. (1999)	Overall series mean ± 0.25 SD	10	1.05	0.94	
APSA	Colby et al. (1999)	Overall series mean ± 0.25 SD	10	1.70	0.94	
MLSA	Wikstrom et al. (2005)	Overall series mean (0–3 s window) ± 0.25 SD	3	1.70		0.80
	Colby et al. (1999)	Overall series mean ± 0.25 SD	10	1.25	0.87	
APTOP	Wikstrom et al. (2005)	Overall series mean (0–3 s window) ± 0.25 SD	3	1.10		0.66
	Ross and Guskiewicz (2004)	Minimal range for 10–15 s or 15–20 s window	3	2.80	0.61	
	Ross et al. (2005)	1.49%BW*	7	1.54	0.79	
MLTOP	DiStefano et al. (2010)	Average range of variation (8–9 s window) + 3 SD	3	2.41	0.37	
	Ross and Guskiewicz (2004)	Minimal range for 10–15 s or 15–20 s window	3	2.24	0.80	
	Ross et al. (2005)	1.71%BW*	7	1.90	0.65	
	DiStefano et al. (2010)	Average range of variation (8–9 s window) + 3 SD	3	1.40	0.78	

* Regards the normalized reference variable, a threshold based on a mean range of variation during single leg stance + 3 SD (Ross et al., 2005).

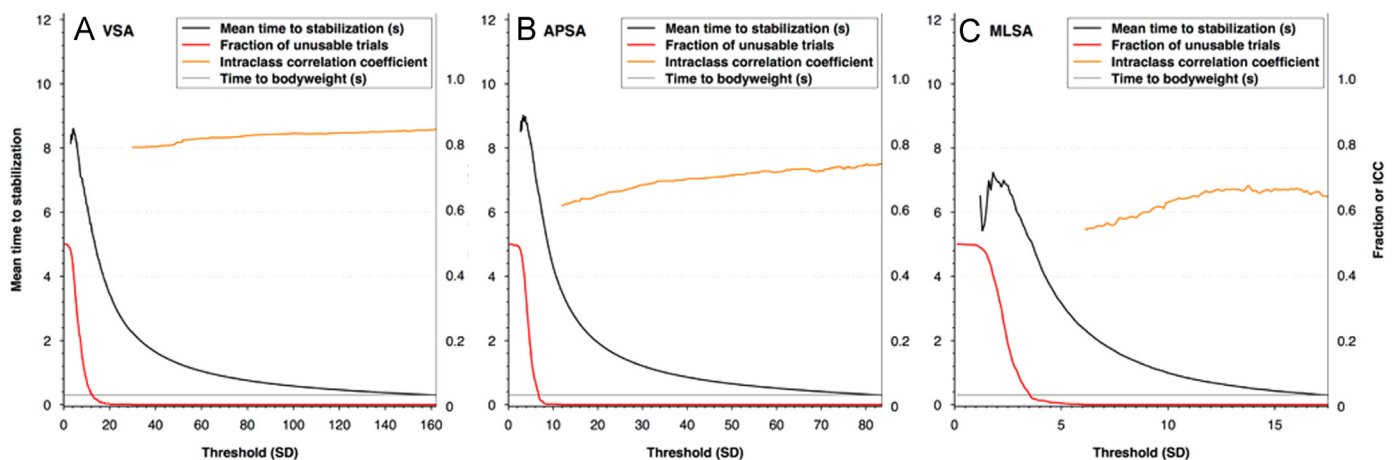


Fig. 5. Mean TTS, fraction of unusable trials, ICC values and mean TTBW with regard to the SA signals (A. Vertical, B. Anteroposterior, and C. Mediolateral). Threshold magnitude along the x-axis differs per panel.

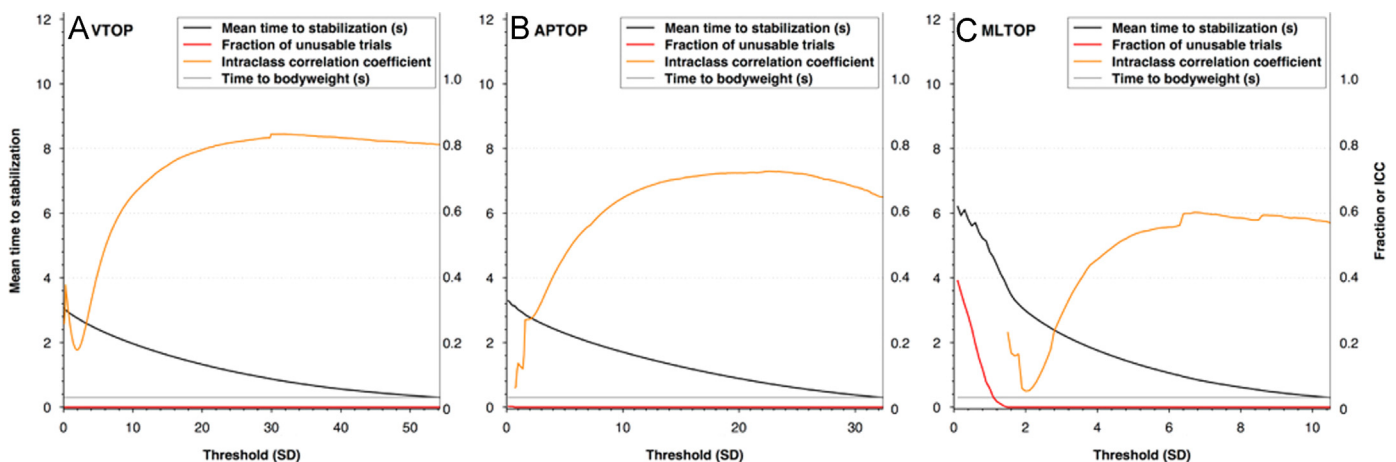


Fig. 6. Mean TTS, fraction of unusable trials, ICC values and mean TTBW with regard to the TOP signals (A. Vertical, B. Anteroposterior, and C. Mediolateral). Threshold magnitude along the x-axis differs per panel.

injury prediction or evaluation, but not a proof. Furthermore, the current sample of elite youth soccer players is very specific; therefore the current results might not be generalizable to other populations.

In conclusion, none of the processed signal/threshold combinations yielded a narrow band of threshold values with optimum reliability. Regardless of threshold, RAW and RMS method are not reliable enough to differentiate between elite youth soccer players,

following a single leg drop jump landing. In contrast, the SA and TOP methods yielded sufficiently reliable TTS values, provided that thresholds are not set too low.

Conflict of interest statement

None of the authors have any personal relationships with other people and organizations that could inappropriately influence or bias this research or the results presented within this manuscript.

Acknowledgments

None.

References

- Brown, C.N., Ross, S.E., Mynark, R., Guskiewicz, K.M., 2004. Assessing functional ankle instability with joint position sense, time to stabilization, and electromyography. *J. Sport Rehabil.* 13, 122–134.
- Colby, S.M., Hintermeister, R.A., Torry, M.R., Steadman, J.R., 1999. Lower limb stability with ACL impairment. *J. Orthop. Sports Phys. Ther.* 29, 444–451.
- DiStefano, L.J., Padua, D.A., Blackburn, J.T., Garrett, W.E., Guskiewicz, K.M., Marshall, S.W., 2010. Integrated injury prevention program improves balance and vertical jump height in children. *J. Strength Cond. Res.* 24, 332–342.
- Ebben, W.P., Vanderzanden, T., Wurm, B.J., Petushek, E.J., 2010. Evaluating plyometric exercises using time to stabilization. *J. Strength Cond. Res.* 24, 300–306.
- Flanagan, E.P., Ebben, W.P., Jensen, R.L., 2008. Reliability of the reactive strength index and time to stabilization during depth jumps. *J. Strength Cond. Res.* 22, 1677–1682.
- Fransz, D.P., Huurnink, A., de Boode, V.A., Kingma, I., van Dieën, J.H., 2015. Time to stabilization in single leg drop jump landings: an examination of calculation methods and assessment of differences in sample rate, filter settings and trial length on outcome values. *Gait Posture* 41, 63–69.
- Fransz, D.P., Huurnink, A., Kingma, I., van Dieën, J.H., 2014. How does postural stability following a single leg drop jump landing task relate to postural stability during a single leg stance balance task? *J. Biomech.* 47, 3248–3253.
- Fransz, D.P., Huurnink, A., Kingma, I., Verhagen, E.A.L.M., van Dieën, J.H., 2013. A systematic review and meta-analysis of dynamic tests and related force plate parameters used to evaluate neuromusculoskeletal function in foot and ankle pathology. *Clin. Biomech.* 28, 591–601.
- Huurnink, A., Fransz, D.P., Kingma, I., van Dieën, J.H., 2013. Comparison of a laboratory grade force platform with a Nintendo Wii Balance Board on measurement of postural control in single-leg stance balance tasks. *J. Biomech.* 46, 1392–1395.
- Ross, S.E., Guskiewicz, K.M., 2003. Time to stabilization: a method for analyzing dynamic postural stability. *Athl. Ther. Today* 8, 37–39.
- Ross, S.E., Guskiewicz, K.M., Yu, B., 2005. Single-leg jump-landing stabilization times in subjects with functionally unstable ankles. *J. Athl. Train.* 40, 298–304.
- Shaw, M.Y., Gribble, P.A., Frye, J.L., 2008. Ankle bracing, fatigue, and time to stabilization in collegiate volleyball athletes. *J. Athl. Train.* 43, 164–171.
- Shrout, P.E., 1998. Measurement reliability and agreement in psychiatry. *Stat. Methods Med. Res.* 7, 301–317.
- Svensson, M., Drust, B., 2005. Testing soccer players. *J. Sports Sci.* 23, 601–618.
- Tulloch, E., Phillips, C., Sole, G., Carman, A., Abbott, J.H., 2012. DMA Clinical Pilates Directional-Bias Assessment: Reliability and Predictive Validity. *J. Orthop. Sports Phys. Ther.* 42, 676–687.
- de Vet, H.C.W., Terwee, C.B., Knol, D.L., Bouter, L.M., 2006. When to use agreement versus reliability measures. *J. Clin. Epidemiol.* 59, 1033–1039.
- Wikstrom, E.A., Tillman, M.D., Smith, A.N., Borsa, P.A., 2005. A new force-plate technology measure of dynamic postural stability: The dynamic postural stability index. *J. Athl. Train.* 40, 305–309.
- Witchalls, J., Blanch, P., Waddington, G., Adams, R., 2012. Intrinsic functional deficits associated with increased risk of ankle injuries: a systematic review with meta-analysis. *Br. J. Sport. Med.* 46, 515–523.